Online Information about Cancer Clinical Trials: Evaluating the Web Sites of Comprehensive Cancer Centers

Valerie Monaco, PhD, MHCI[1,2] and Sandra K. Krills, BS[2]

[1] Center for Biomedical Informatics, University of Pittsburgh, Pittsburgh PA
[2] Benedum Oncology Informatics Center, University of Pittsburgh Cancer Institute, Pittsburgh PA

ABSTRACT

The objective of this study was to examine the information provided on comprehensive cancer centers' Web sites regarding clinical trials. Thirty-nine Web sites were visually inspected for four categories of variables: navigation to the clinical trial information, search functionality provided to the visitor, information content provided about trials, and the reading level of the information provided. Results indicated that for those Web sites that provided information about clinical trials, the content was often limited and trial descriptions were written at a college reading level. This study suggests that these Web sites are not yet adequately designed to serve as a successful aid for increased trial accrual. The design of future online clinical trial information should be guided by data from consumer health informatics research.

INTRODUCTION

Over the past several decades, new and more effective treatments have been made available to cancer patients. These breakthrough treatments are a direct result of cancer clinical trials that have been conducted to evaluate the efficacy of new treatment approaches. Despite the promise of additional breakthroughs, only a small percentage (estimated at 2-3%) of eligible adult cancer patients enroll in clinical trials.[1] This low rate of accrual represents a bottleneck for researchers attempting to identify new ways to treat cancer.

Many studies have been conducted to examine factors related to these low enrollment rates.[2,3,4] For example, a recent Harris Poll of cancer patients revealed that many patients may have a lack of knowledge regarding clinical trials or may hold misconceptions.[5] A large majority of patients surveyed (85%) indicated that they didn't know that enrollment in a clinical trial was an option for them. Those that did know about clinical trials often shied away from pursuing enrollment due to fears of getting a placebo treatment, receiving substandard treatment regimens, and not having insurance coverage for the trial. Clearly, education of and information access for patients must be critical components to any plan addressing low enrollment rates in cancer clinical trials.

One of the most effective channels for providing health-related information access is proving to be the Internet. In 2002, 62% of Internet users reported that they had used the Internet to search for health-related information.[6] Although there are many sites on the Web that provide health information, hospital Web sites are becoming a destination for more Internet users. A recent survey revealed that hospital site visitors more than tripled from 2001 to 2002.[7] Hospital Web sites represent an important access point for information about cancer clinical trials because no comprehensive database exists of cancer clinical trials. Hospital Web sites may represent the only place that information about a particular trial exists on the Internet.

The scope of what is known regarding online information about cancer clinical trials is mostly limited to descriptions of development efforts that led to several online resources.[8,9] To date, no content review of online information about cancer clinical trials has been conducted. The goal of this research study was to examine clinical trial information on the Web sites of National Cancer Institute (NCI) designated comprehensive cancer centers (CCC's). According to the NCI, these centers are "encouraged to initiate and conduct early phase, innovative clinical trials and to participate in the NCI's cooperative group system by providing leadership and accruing patients to trials."[10]

Reflecting the adage, "knowledge is power", this study assessed how comprehensive cancer centers are using the Internet to inform patients about trials and, in turn, empowering patients to make decisions regarding clinical trials. Specifically, we examined four categories of Web site variables: navigation to the clinical trial information, search functionality

provided to the visitor, information content provided about trials, and the reading level of the information provided.

METHODS

The URL addresses for the NCI CCC's Web sites (N=39) were obtained from the NCI's Cancer Centers Web site[11] and accessed between February and March 2003. Except for redirects, the first page returned by the provided URL was considered the home page for that site. Web sites were visually inspected by one of the authors (SKK) for the following categories of variables:

Navigation to the Clinical Trials Information: We noted whether there was a link on the home page that said "Clinical Trials" (either immediately visible or revealed by a rollover effect). The number of mouse clicks to the first labeled clinical trials page was recorded. If no link on the home page suggested how to get to clinical trials information, the site index and search feature, in that order, were accessed using the string "clinical trials."

Search Capabilities: We examined the clinical trials section of the Web sites for available searching functions and features. Specifically, we explored: Did the Web site allow users to search for clinical trials by cancer type? Did the Web site use both lay and medical terms for cancer types? This was operationalized by noting whether the term "skin cancer" was used along with or in place of "melanoma." We also noted how many trials were returned for breast cancer, prostate cancer, and lung cancer and the manner in which the returned trials were organized.

Information Content: Prior to data collection, we brainstormed a list of possible items that could be included about a particular clinical trial. The final list contained 41 items. The items were inspired by elements observed on other clinical trial Web sites (e.g., government Web sites devoted to clinical trials), research findings about patients' attitudes and concerns about clinical trials (e.g., fears of getting a placebo), as well as features on commercial Web sites (e.g., "ask the research team a question via e-mail" is similar to a feature on e-Bay). The information provided about several trials on each site was read before noting the presence of these information items. In addition, we coded whether the site contained any brief explanatory text regarding clinical trials and/or a more comprehensive FAQ-like document.

Readability of Trial Description: The information items provided to users were presented in different formats on different sites. Some sites simply provided a label and value pairing (see Figure 1.a). Other sites provided information using a sentence or list format (see Figure 1.b). For those sites that provided trial information in sentence/list format, we analyzed the reading level of the provided information. For these analyses, we attempted to collect up to nine trial descriptions from each Web site (3 breast cancer trials, 3 lung cancer trials, and 3 prostate cancer trials). Most trial descriptions analyzed were from treatment-related protocols. A total of 111 descriptions were collected and analyzed from 17 sites.

(a)

Protocol Number	Title	Principal Investigator
ECOG N9841	A Randomized Phase III Equivalence Trial of Irinotecan (CPT-11) Versus Oxaliplatin (OXAL) / 5 Fluorouracil (5-FU) / Leucovorin (CF) in Patients with Advanced Colorectal Carcinoma Previously Treated with FU	Smith

(b)

Purpose : FK866 is an experimental drug which has shown encouraging results in the laboratory in fighting cancer. It is thought to interfere with the metabolism of a cancer cell, resulting in cell death. FK866 has not yet been studied in patients.

Figure 1. (a) Example of label/value pair trial display. (b) Example of sentence/list trial description display.

All contiguous information regarding a particular trial was copied from the Web site into a Microsoft Word 2000 document. The reading level of the material was then analyzed using the Microsoft algorithm that computes the Flesch-Kincaid reading level. The Flesch-Kincaid reading level[12] is based on the average number of words per sentence and the average number of syllables per word. The Microsoft implementation of this formula caps the highest reading level at the 12th grade. In order to determine reading levels for passages that exceeded the 12th grade level, we created a Visual Basic macro that counted syllables, words, and sentences. This macro was benchmarked against the Microsoft algorithm using a pre-identified group of 15 passages written below the 12th grade level. The correlation coefficient between the two methods for computing the Flesch-Kincaid scores was 0.98. The reading level of the trial description was computed using both methods.

All Web sites were viewed using Internet Explorer 6.0 on a Compaq computer (Microsoft 2000 OS) located at the Benedum Oncology Informatics Center. All data entry was recorded into an Excel 2000 spreadsheet during the visual inspection and during the readability analyses. The study was approved by the University of Pittsburgh Institutional Review Board (IRB).

RESULTS

Navigation to the Clinical Trials Information: Twenty-nine (74.3%) of the centers provided a "clinical trials" link on their home page. The main page for clinical trials was accessible by one mouse click on 29 sites. For four (10.3%) sites, the information was accessible in two clicks, and for the remaining six (15.4%) sites, the search feature on the site needed to be used to find the clinical trials section.

Search Capabilities: Seven center Web sites could not be searched further for clinical trial information. Of those seven, three sites had text indicating that their clinical trials section was under development, two only provided links to external sites containing information about clinical trials (e.g., www.cancer.gov), and on two sites information about clinical trials could not be found using the navigation methods described above. For the remaining analyses, the number of sites is 32 (except for the FAQ and explanatory text variables which was provided on some of the sites that did not have trial-specific information).

Twenty-five (78.1%) of the remaining sites allowed users to search for trials by cancer type at some point during the selection process. Thirty sites (93.4%) had access to melanoma trials, but only 11 (36.7%) used the term "skin" or "skin cancer" in their selection options.

Selecting breast cancer trials returned an average of 18.7 trials (range 3-46). Lung cancer searches returned an average of 11.1 trials (range 1-40). Prostate cancer searches returned an average of 10.1 trials (range 1-27). (Note: the numbers of returned trials for lung and prostate cancer are conservative because oftentimes these trials were embedded in lists of thoracic and genitourinary trials, respectively. The lung and prostate cancer trials were extracted from these longer lists.)

For 16 (50%) sites, the order in which the returned trials were presented could not be determined. Nine (28.1%) organized the trial listings by a protocol identification number (presumably assigned by that institution's IRB). The remaining sites organized the trials by organ system (n=3), cancer stage (n=1), visual decision tree (n=1), or a combination of variables (n=2).

Information Content: Of the original 41 content items that we examined trial listings for, 10 items appeared on 4 or more sites (i.e., more than 10% of the sites). Table 1 summarizes these findings.

Table 1. Frequency of Trial Information Content that appeared on four or more web sites.

Information Content	# of Sites (%)
Title of Study	32 (100.0)
Protocol ID	30 (93.8)
Contact Phone Number	30 (93.8)
Contact E-mail	25 (78.1)
Principal Investigator	19 (59.4)
Eligibility Requirements	17 (53.1)
Purpose of Study	12 (37.5)
Date of Trial Listing	10 (31.3)
Treatment Details	5 (15.6)
Intended Audience for Trial Description	5 (15.6)

In addition to trial details, 22 sites (56.4%) provided brief explanatory text about clinical trials (e.g., "Clinical trials are research studies that are conducted in order to evaluate potential new treatments for cancer") and 14 sites (35.9%) provided a more comprehensive FAQ-like document about clinical trials. (Note: these elements could be present without actual trial listing and therefore the percentage reflects all 39 Web sites.)

Readability of Trial Description: The average grade level of the 111 narratives as calculated by the Microsoft 2000 Flesch-Kincaid formula was the 11th grade (11.43 years of school). The Visual Basic macro that permitted reading levels to be recorded above the 12th grade returned an average reading level of 14.83 years of school. Because the Flesch-Kincaid formula is determined in part by the average syllables per word, we were concerned that the drug names included in the description (which are often multi-syllabic) may have artificially inflated the reading levels. Consequently, a subsequent analysis was done with the drug names removed from a separate sample of trial descriptions. This analysis revealed that removing the drug name did lower the reading levels but the average reading level returned by the Microsoft 2000 Flesch-Kincaid formula remained at the 11th grade level (11.26 years of school).

DISCUSSION

The public is increasingly turning to the Internet for health-related information. This suggests that the web sites of the NCI CCC's represent an important avenue for presenting information to patients about available cancer clinical trials. Yet despite continual calls for identifying ways to improve accrual rates, currently available online information about clinical trials appears to have been assembled and presented without that goal in mind.

The data from these 39 Web sites suggests three likely scenarios for visitors. In the first scenario, visitors are able to retrieve a list of trials and can access trial descriptions one by one, but find that the descriptions are written at a much higher level than they can understand. The findings of this study suggest that the average trial description is written at a college reading level. The average adult reading level in the United States is at an 8th or 9th grade level.[13] Although many complex medical terms can not be expressed using simpler terms, overall trial descriptions must be presented at a more accessible level.

In the second scenario, visitors are able to retrieve a list of trials by cancer type, but are given only minimal information about the trial (e.g., protocol ID, title of study, principal investigator's name and a contact phone number). Although not analyzed separately in this study, the titles of clinical trials are often difficult to understand given that they can be quite lengthy and contain a number of scientific and technical terms. It is unlikely that a visitor reading a dozen scientific titles would be able to begin a decision-making process regarding enrollment in a trial. In the last scenario, visitors may not be able to locate or access any clinical trial information.

The field of consumer health informatics is uniquely positioned to address the design of online clinical trial information. We need a better understanding of patients' information needs regarding clinical trials, and we also need a better understanding of how best to present that information. Only with this research will we be able to effectively deliver online clinical trials information.

This study's limitations included a lack of inter-rater reliability. In addition, the formatting of the trial descriptions taken from the Web was not always optimal for computing readability scores (e.g., some listed items appeared without punctuation). Lastly, because we did not study actual users while they navigated the sites, we can only speculate on the difficulties that they may encounter when attempting to gather information about clinical trials.

In sum, future efforts at increasing accrual rates to trials should not overlook the possibilities that enhanced Web sites may yield. Data-driven design, informed by consumer health informatics research, will provide patients and caregivers with optimal online clinical trial information resources.

REFERENCES

1. Tejeda HA, Green SB, Trimble EL, et al. Representation of African-Americans, Hispanics, and Whites in National Cancer Institute cancer treatment trials. J Natl Cancer Inst. 1996 Jun 19;88(12):812-6.
2. Gotay CC. Accrual to cancer clinical trials: Directions from the research literature. Soc Sci Med. 1991;33(5):569-77.
3. Lara PN, Higdon R, Lim N, et al. Prospective evaluation of cancer clinical trial accrual patterns: Identifying potential barriers to enrollment. J Clin Oncol. 2001 Mar 15;19(6):1728-33.

4. Ellis PM, Butow PN, Tattersall MHN, Dunn SM, Houssami N. Randomized clinical trials in oncology: Understanding and attitudes predict willingness to participate. J Clin Oncol. 2001 Aug 1;19(15):3554-61.
5. Misconceptions and low awareness slow recruitment for cancer clinical trials. http://www.harrisinteractive.com/news/allnewsbydate.asp?NewsID=213 (1 Mar 2003).
6. Fox S, Rainie L. Vital decisions: How Internet users decide what information to trust when they or their loved ones are sick. http://www.pewinternet.org/reports/reports.asp?Report=59&Section=ReportLevel1&Field=Level1ID&ID=262 (1 Mar 2003).
7. Cybercitizen ® Health v2.0: Hospital and Health Plan Site Research Module. http://www.manhattanresearch.com/Hospital%20and%20Health%20Plan%20Module.pdf (1 Mar 2003).
8. McCray AT, Ide NC. Design and implementation of a national clinical trials registry. J Am Med Inform Assoc. 2000 May-Jun;7(3):313-23.
9. Stahl DC, Stahl CL, Niland JC. Evaluating the impact of *Clinical Trials On-Line* on clinical trial awareness and accrual. Proceedings of the 33rd Hawaii International Conference on System Sciences; 2000 Jan 4-7; Maui, Hawaii.
10. Cancer Center Designations. http://www3.cancer.gov/cancercenters/ccsg_comp_pt1_6to7.html#6cancer (1 Mar 2003).
11. NCI-designated Cancer Centers. http://www3.cancer.gov/cancercenters/centerslist.html (Feb/Mar 2003).
12. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy enlisted personnel. Research Branch report 8-75. Memphis: Naval Air Station,1975.
13. What is Health Literacy? http://www.fhidc.com/pfizer/whatis_stats.html (1 Mar 2003).

ACKNOWLEDGEMENTS

The authors wish to thank Beth Simon, RN, MSN, Rayna Robb, BA, Deborah Polk, PhD, and Adrienne Weiss, BA for their assistance with this project. Funding support for this study was provided by the Benedum Foundation and the Emma Clyde Hodge Foundation.